A Simple Multiple Linear Regression Test for Differential Effects of a Given Independent Variable on Several Dependent Measures

Jerry A. Colliver
Steven J. Verhulst
Paul Kolm
Southern Illinois University
School of Medicine

ABSTRACT

Multiple linear regression may be used to determine whether an independent variable of interest has a differential effect on two or more dependent variables. The initial step involves the separate standardization of each dependent variable. The values of the standardized dependent variables are pooled and treated for purposes of the analysis as constituting a single dependent variable. A within subjects independent variable is created and the levels of the variable are used to denote the different dependent variables. The data are analyzed with a split-plot analysis of variance for which the independent variable of interest is the between groups factor and the independent variable which distinguishes the dependent variables is the within subjects factor. The test of the interaction of these two factors provides a statistical determination of whether the independent variable of interest has a differential effect on the two or more dependent variables.

A problem we have encounted on several occassions can be dealt with easily by using an interesting "twist" on multiple linear regression procedures. The problem involves the determination of whether a given independent variable has different effects on several dependent measures. For example, most recently, we were asked to determine if the dosage of a given drug administered to animals injected with tumor cells had different effects on tumor size and body weight. To make this determination, we separately standardized each of the two dependent variables, tumor size and body weight, pooled these standardized values, and treated the two standardized variables as if they constituted one dependent measure. The two standardized dependent variables were distinguished via a within subjects, independent variable (called Outcome Measure), which we created for the purpose. This within subjects, independent variable had two levels which denoted the two standardized dependent variables, respectively. A split-plot analysis of variance (ANOVA) was performed and the test of the Dosage X Outcome Measure interaction provided a simple test of whether Dosage had different effects on the two outcome measures, tumor size and body weight.

PROCEDURE

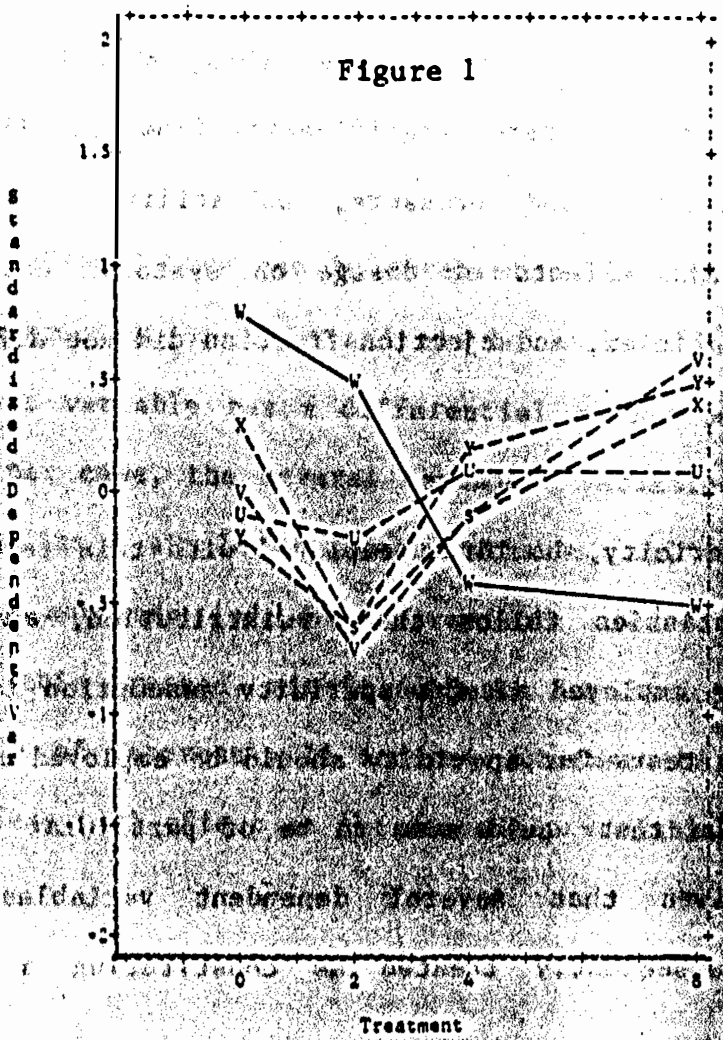
The procedure can be illustrated with a set of simulated data used to stimulate "solutions" for discussion purposes at a recent Multiple Linear Regression Special Interest Group session (Leitner, 1986). (See Appendix A.) Data were generated for n = 30 hypothetical subjects on five continuous variables (Y, X, U, V and W) and three dummy variables (D1, D2 and D3). For the purpose of illustrating the procedure, the five continuous variables were regarded as dependent variables. Each was standardized, and the five standardized variables were subsequently treated for purposes of the analysis as representing one dependent variable. The five variables were distinguished by considering each variable as if it represented one level of an artificially created independent variable, Outcome Measure.

The three dummy variables, D1, D2, and D3, were treated as if they represented one independent variable called Treatment with levels represented by the binary code expressed by the three dummies. Using this procedure the independent variable was found to have four levels represented by the binary codes, 000, 010, 100, and 111. Thus, the four levels of the Treatment independent variable were 0, 2, 4, and 8.

A 4 X 5 split-plot analysis of variance with one between subjects variable (Treatment with four levels, 0, 2, 4, and 8) and one within subjects variable (Outcome Measure with five levels, Y, X, U, V, and W) was performed on the simulated data. Treatment represented the independent variable of interest and Outcome Measure represented the independent variable used to distinguish the five standardized dependent variables.

RESULTS

The results showed a significant Treatment X Outcome Measure interaction, indicating that Treatment had different effects on the different outcome measures, $F(12,104) = 2.21$; $p = 0.0448$. Simple interaction effects tests showed that the effect of Treatment on the dependent variable W differed significantly from the effects of Treatment on the other four dependent variables, Y, X, U, and V, and that the effects of Treatment on the four dependent variables, Y, X, U, and V, did not differ significantly. A graph of the relationship between Treatment and the five dependent variables is presented in Figure 1, which shows that variable W decreased from Treatment level 0 to level 2 to level 4 and remained fairly stable from level 4 to level 8. Variables Y, X, U, and V decreased from level 0 to level 2, increased from level 2 to level 4 to level 8.



DISCUSSION

The results showed that the independent variable, Treatment, had significantly different effects on the five dependent variables, Y, X, U, V, and W. To give substance to this example, suppose that the Treatment independent variable with four levels represented the dosage of some drug such as ethanol, epinepherine, streptokinease, etc. and that the four dosages were 0, 2, 4, and 8 units. Further suppose that the five dependent variables were as follows: Y, systolic blood pressure; X, diastolic blood pressures; U, pulsatility index; V, ejection fraction; and W, heart rate. The research hypothesis, then, would state that drug dosage has a differential effect on the five dependent variables, and the null hypothesis would be H_0: σ^2(interaction) = σ^2(error). Our results, then, showed that the effect of drug dosage on heart rate differed significantly from the effects of dosage on systolic and diastolic blood pressure, pulsatility index, and ejection fraction but that the effects of dosage on systolic and diastolic blood pressure, pulsatility index, and ejection fraction did not differ significantly from one another.

The test for spericity should be employed with this test to determine if the computed F statistics follow the F distribution, and an appropriate adjustment should be employed if the spericity assumption is violated (Kirk, 1982). Although the tests for spericity should be employed routinely with any split-plot ANOVA, the test would seem to be of particular importance in the present context given that several dependent variables are separately standardized and subsequently treated as constituting a single dependent variable.

The reader will undoubtedly notice the similarity between the procedure outlined here and the more commonly known profile analysis (Morrison, 1967). The difference in emphasis and orientation between this procedure and profile analysis, however, would seem to warrant separate consideration of the procedure described here. Profile analysis focuses on the comparison of profiles of means of several variables for two or more groups. The typical example involves the comparison of profiles of means on psychological tests in a test battery for groups of patients with different psychiatric diagnoses. The typical graphic representation depicts a profile of test (dependent variable) means plotted separately for each group. The procedure outlined here, on the other hand, involves the comparison of the effects of an independent variable on several dependent variables, with a graphic representation that depicts the effect of the independent variable on each dependent variable separately (see Figure 1).

The procedure outlined here can be extended to designs with more than one between groups, independent variable and can be used to determine if a within subjects independent variable has a differential effect on several dependent variables. In either case, the several dependent variables are standardized, treated as constituting a single dependent variable, and distinguished by the levels of a within subjects independent variable created for that purpose. The interaction of this created, within subjects independent variable and the independent variable of interest will indicate whether the latter independent variable has a differential effect on the dependent variables.

REFERENCES

Kirk, R.E. (1982). Experimental design: Procedures for the Behavioural Sciences (2nd ed.). Belmont, CA: Brooks/Cole.

Leitner, D.W. (1986). Data set for you to analyze. Call for papers for Multiple Linear Regression Special Interest Group, American Education Research Association, San Francisco, CA.

Morrison, D.F. (1967). Multivariate statistical methods. New York: McGraw-Hill.

OBS	Y	X	U	V	W	D1	D2	D3
1	54	47	49	62	41	0	1	0
2	42	55	64	56	66	1	0	0
3	64	61	47	81	49	0	1	0
4	48	45	63	55	46	0	0	0
5	5	21	93	31	62	1	1	1
6	42	46	11	50	53	1	0	0
7	40	55	14	54	67	1	1	1
8	62	55	26	74	44	1	1	1
9	45	56	13	59	63	1	1	1
10	47	43	52	52	44	1	0	0
11	61	69	96	83	63	0	1	0
12	62	69	11	84	62	1	0	0
13	54	41	30	58	33	0	0	0
14	47	47	83	55	49	1	0	0
15	48	38	69	51	36	0	1	0
16	87	78	49	47	47	0	1	0
17	47	51	31	58	55	1	0	0
18	73	49	35	40	70	1	1	1
19	49	49	73	58	50	0	0	0
20	40	43	92	46	53	1	0	0
21	54	44	47	50	37	0	0	0
22	52	49	54	61	47	0	0	0
23	48	47	70	56	48	0	1	0
24	40	45	10	65	34	0	1	0
25	37	43	96	43	56	1	1	1
26	39	40	26	43	48	1	1	1
27	46	47	56	54	49	1	1	1
28	62	58	48	76	48	0	1	0
29	46	44	53	52	47	1	1	1
30	35	40	63	39	54	1	1	1

Appendix A. Simulated Data from Multiple Linear Regression Special Interest Group Session (Leitner, 1986).